Shansan Gong

hisansas@gmail.com · (+86) 131-6207-6936 · WebPage: https://summmeer.github.io/

Education

Shanghai Jiao Tong University, Bachelor, EE, GPA: 3.84/4.3 Top 5% 2015.09 – 2019.06

Shanghai Jiao Tong University, Master, CS, GPA: 3.78/4 Top 10% 2019.09 – 2022.03

Experience

My research interests include Machine Learning, Natural Language Processing and Information Retrieval.

Diffusion Language Model Research Program 2022.06 – Present

Propose a diffusion model DiffuSeq, designed for Seq2Seq text generation tasks. Upon extensive evaluation over a wide range of Seq2Seq tasks, we find DiffuSeq achieving comparable or even better performance than six established baselines, including a state-of-the-art model that is based on pre-trained language models.

Dynamic Multi-Domain Product Categorization Industry Program 2021.12 – 2022.03

For e-commerce platforms, when trying to categorize the products, there are multiple domain-specific category taxonomies and each of them evolves dynamically over time.

- We propose a retrieval and rerank framework. Considering that pure vector matching may fall for the surface form of text, we further leverage the universal concept of "knowledge" across different domains to complement textual semantics. Experiment results verify the knowledge integration cross domains.

Session-based News Recommendation Research Program 2020.06 – 2021.09

News recommendation for anonymous users is useful on many news portals. It is challenging because both the lifespan of news articles and the duration of user visits are short. Under the session-based scenario, we model the news recommendation task as the Next-Item Prediction task.

- We leverage the positive/negative and neutral implicit feedback of the user to figure out to what extend the user likes/dislikes the article, which better tackles the user cold-start problem: we can make an accurate prediction about anonymous users earlier in the session.
- By proper representation of the content and temporal information in the sessions, our approach better handles the article cold-start problem and mitigates the dilemma between diversity and accuracy.

Internship

Microsoft STCA NLP Engineer 2021.06 – 2021.09

In order to optimize the search result, it is necessary to judge whether the intents of different search Query1 and Query2 are consistent, which is formulated as a binary classification task.

- Based on the cross-language pre-training model InfoXLM, introduce PostWeb information (ie, text information such as the title, link, and abstract of the Query search result) as an additional feature of the Query.

Publications

[1] *First author. Positive, Negative and Neutral: Modeling Implicit Feedback in Session-based News Recommendation (Accepted by SIGIR 2022)*

[2] *First author. Enhanced Semantic Space: Integrate Concepts to Unify Dynamic Multi-Domain Product Categorization (AAAI 2023 Under Phrase II Review)*

[3] *First author. Sequence to Sequence Text Generation with Diffusion Models (ICLR 2023 Under Review)*

Honors and Awards

2nd Prize, China Post-Graduate Mathematical Contest in Modeling Nov. 2020

Shanghai Jiao Tong University Wish Scholarship 2021

Shanghai Jiao Tong University Shenzhen Stock Exchange Scholarship 2020